

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: MCGILL UNIVERSITY
- (ii) TITLE OF INVENTION: THE C. ELEGANS gro-1 GENE
- (iii) NUMBER OF SEQUENCES: 62
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: SWABEY OGILVY RENAULT
  - (B) STREET: 1981 McGill College Avenue - Suite 1600
  - (C) CITY: Montréal
  - (D) STATE: QC
  - (E) COUNTRY: Canada
  - (F) ZIP: H3A 2Y3
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: Windows
  - (D) SOFTWARE: FastSEQ for Windows Version 2.0b
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: PCT/CA98/00803
  - (B) FILING DATE: 20-AUG-1998
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: CA 2,210,251
  - (B) FILING DATE: 25-AUG-1997
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Côté, France
  - (B) REGISTRATION NUMBER: 4166
  - (C) REFERENCE/DOCKET NUMBER: 1770-179"US" FC/gc
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 514 845-7126
  - (B) TELEFAX: 514 288-8389
  - (C) TELEX:

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 14458 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAAAATTTG	CTAAGATGAA	GCGCCGGCTT	GTTACATTGC	TTTTCAGAGT	CGATTGGTTC	60
AAAATTGTCA	ATTTTATCCA	AAATAGAGTG	CATTGTGTGT	ACAATAACTA	AAGAATCATC	120
CATATCTGGT	CCAACACAAC	ATTGATGGAA	TACTGGATCA	ATTGTCTAAA	AAAATATCAA	180
TAGAATAATG	AAACATTTTC	AGAATTCATT	ACCGTCAATG	TCAGATAGTC	ATTCCTTGAG	240
TATTTTGTGG	ATGCTTTGAA	AATTCTTCGC	TGGGCCATAT	CTGTTGGATA	ATCTGAAAAA	300
CGCAATAAAT	TTCATCGAAA	ATGCCCTATTA	AATTGAATTA	CCTTCTTCTT	CATCATTTC	360
TAACAATTCA	TGCTCTTTTT	GTGCTTGACT	TGTGACCAAT	TCTTTAAATT	CAATTAAATC	420
GTCAATATCC	TTTTGTACTA	AATCCATCTT	GATATTCAAT	ATATCTTTGT	CAGTATAGTA	480
TTCAGCGTAT	CTGAAATTTT	GAATTTATTT	TTCTAATTCC	CAAGAAAAAT	AATTAATAAG	540
AATACCTTAA	CGAATTATTA	TCCAATATAT	CATCATTGTC	CACATCTGGA	AGACGCTGAG	600
GAAGTGTGTT	AGCAGCTTGG	AGGTAGTCGT	CATCGTCTCT	GGAAATTGTT	ATTTTCAATT	660
TCAAAAAAAA	AACTTTACTT	ACGAAATATA	CTCATTTGAT	GCAATCCACG	GATCAAAACG	720
ACGTCTTTGC	ATCTTTGAAT	CATTTTCCGC	ATGGCACCGC	ATCACTTCTT	TCTTATGATT	780
ATTTTCTAAC	GTTTGTGAAA	ATTCGACGTG	CTCTTCACAA	CGGCCGCCAT	GTTTCGCAAG	840
TTCTTCTTTT	GATCGTATCT	AAAATTTTAA	ATTTGAAAAA	AAGCTTACTA	TCAAATTTTC	900
GTATTTTTTT	TCACCTGCTT	ACACCGAACA	AGCGTTCGAT	ACGAAGCATA	ATTACATTGT	960
CCATACTTAT	TTTTGTCTGA	TTCATTGGCA	ACAAGACGGA	ATCGTGTTC	AGGTGCAACT	1020
ATATATTGAG	CAGGAGGACG	AGTTGTTTGT	TTCATGCTGC	TTAAAAATAA	AAATGGAAAA	1080
TTGAGTCAAA	AAGTTGAGAT	AAAACAAATT	AAAACAATTT	TCTGAAAAAT	AAACAACCTGA	1140
AATTTGAAGT	AATAAACAAAC	ACGCGAAAAAC	GTTATTTCGG	AGCATCGTTT	GAGAAGTAAA	1200
ACTTTTTTTT	GGCGCACCCCT	TGTGCGCAGT	TTTTATCTTC	TCTTTTAATT	TAATTTTCAA	1260
GCTAAATCTT	TCTTTTTTAA	CTTTGAATAA	ATATTTAAAT	ATTCAGAATG	CACCAATAAA	1320
CCTGGAACAA	AATCGATAAT	GTTCCGCAAG	CTTGTTCTT	CTGGGTCACT	ATGGAAGCCG	1380
AAAAATCCGC	ATTCTTTGGA	ATACCTCAAA	TATTTACAAG	GAGTGCTCAC	AAAAAATGAG	1440
AAAGTTACGG	AAAACAATAA	GAAAATATTA	GTAGAAGCAT	TACGAGCTAT	CGCAGAAATT	1500
CTCATTTGGG	GCGATCAGAA	TGATGCTTCG	GTTTTTGAGT	GAGTTTTTTT	CCAATGTTTT	1560
TTTTCAAATC	TGATGTTGAA	TTTCAGTTTC	TTCTTGAGC	GGCAAATGCT	TCTTTATTTT	1620
TTGAAAATTA	TGGAACAAGG	AAACACACCA	CTAAATGTAC	AATTACTGCA	GACTTTGAAC	1680
ATTTTATTTCG	AAAATATTTCG	ACATGAAACT	TCACTTTGTA	AGTTTTTTTAT	ATGGATTTTC	1740
GCTTAAATTT	GCCAGTTTTC	AGATTTTCCTT	CTAAGTAACA	ATCATGTAAA	CTCGATTATT	1800
TCCCACAAAT	TCGATTTACA	AAATGATGAG	ATCATGGCTT	ACTACATTAG	TTTTCTGAAA	1860
ACTCTTTTCAT	TTAAACTGAA	TCCAGCTACA	ATCCACTTCT	TCTTCAATGA	AACGACTGAA	1920
GAATTTCCAT	GTTTGGTAGA	AGTTTTGAAG	CTTTATAAAT	GGAATGAATC	AATGGTTCTGA	1980
ATTGCTGTTA	GAAATATTCT	TTTAAATATT	GTGAGAGTTC	AAGATGATTTC	AATGATTATT	2040
TTGCTATCA	AGCATACAAA	AGTTAGTAGA	AAATTATTTT	GAAAAGGTGT	ATTTAAGCAA	2100
TAAATATTAC	AGGAATATCT	ATCGGAGTTA	ATAGATTCTC	TAGTTGGTCT	CTCACTTGAA	2160
ATGGACACAT	TTGTACGATC	TGCTGAGAAT	GTGTTAGCTA	ATCGAGAGAG	ATTACGAGGA	2220
AAAGTGGATG	ATTTAATTGA	TTTGATTTCAT	TATATTGGTG	AACTATTGGA	TGTGGAAGCT	2280
GTCGCCGAAA	GTTTATCAAT	TTTAGGTCAG	TTTTACTGCT	GGAAAATCAA	GTTTTTAATG	2340
TTAAATTTTC	AGTAACAACA	CGATACTTAA	GCCCTCTATT	ACTTTCAAGT	ATATCACCAA	2400
GAAGAGATAA	TCATTCACTT	CTACTCACTC	CGATTTCTGC	GTTATTTTTT	TTCTCTGAAT	2460
TTTTATTGGT	GAGTTTTAAC	ATTTAAAATT	ACATTTTCTT	AATTTATTTA	TTTTTCAGAT	2520
AGTTCGTCAC	CATGAAACAA	TATATACATT	TTTATCATCT	TTCTATTTTG	ACACTCAGAA	2580
TACTTTGACG	ACCCATTGGA	TACGTCATAA	TGAGAAATAT	TGCTTAGAAC	CGATTACATT	2640
ATCATCACCA	ACCGGAGAAT	ATGTGAATGA	AGACCAGTAA	GAGCTGAAAT	TTTAAAATTT	2700
TTGCTTTGAA	TATAGTATTT	TCAGCGTATT	TTTCGATTTT	CTACTGGAAG	CATTTGATTTC	2760
CAGTCAAGCA	GACGATTCGA	AGGCATTCTA	TGGATTAATG	CTGATTTATT	CAATGTTTCA	2820
GAATAATGGT	GAGTTTTTAA	AAATTGATTT	GTTAAATTAA	AATTTCCATT	TCCAATAACT	2880
CCTCTTCAGA	CAGTAAGTTT	TCAATGTTGT	AAAGTTCCTG	TTCATCTGTG	ATCGTTTTCT	2940
TCATTTTTTT	AGTTTTGCAT	GAACAGTTTT	CAAATTTTTT	TGATATCATA	CAGTAAATAT	3000
CGTCATCCAG	ATAATTTTCT	ATTTAAAAAA	AATGAATAAA	AAGAGGGCGC	GCAGAAATTG	3060
CCGAAGTAAT	GTAAATTTAA	AGGGACACAT	GCGTAGCTTG	TTGTGTGGGT	CTCGCCGCGC	3120

TTTGTGTTGAT	TTATCTTGTT	TTCTGCTCAA	AGAGCTGTTT	TTATTTTAGC	GTTGAATGCT	3180
TTTTTACCGT	TCTCATCGGC	TTTTTAATAG	GAATATTTAA	AAAAAAAGGT	TTAATAAATC	3240
TTCGTTTTTA	CAAAATCCAT	CTAAGATTTG	CATTTGTGAA	GCTCAACAAG	TAAAGTTTTA	3300
AGTAACATTG	TTTTTTAAAA	AACAATTGAA	CCAAATTTTG	CCGAAACATT	AATAACATGA	3360
CGATACTCTA	TAAATATATC	CTCTTTTCAA	AATAAATTTT	CAAAAAAAT	CCATTTTTCA	3420
GCCGATGTTG	GAGAACTTCT	ATCTGCTGCC	AACCTCCCAG	TGCTCAAAGA	ATCAACGACA	3480
ACTTCATTAG	CTCAACAGAA	TCTTGCTCGT	CTCCGAATAG	CATCTACGTC	TTCCATATCA	3540
AAGCGAACGA	GAGCTATCAC	TGAAATTGGA	GTAGAAGCGA	CCGAGGAAGA	TGAGATTTTT	3600
CATGATGTTT	CTGAAGAACA	AACGTTGGTA	AGTAAATAAA	TCAACATTGA	TTGTTACACA	3660
AACTTTAATA	TTTTTAAATT	TGAAAATTTT	CTTCAAAGTG	CTCAAAAATC	CTGTCGAAAA	3720
TTACAGGAAG	ATCTGGTGGA	TGATGTATTG	TTTGATACTG	AAAATTTCAG	AATCAATGAT	3780
CCAGAAGTGA	GTAGAAAACG	TGCATGTATT	AATTATTAAA	AAAAAAATAT	AGTTTTCCCC	3840
AGTTTTCTTT	GACCTAAAAAC	TCAGCAATTT	CAGCCTAAAA	ACGTGGAGTC	AGAATCTCGT	3900
TCTCGATTTT	AATCTGCTGT	TGATGAGCTT	CCACCTCCGT	CGACTTCTGG	ATGTGATGGT	3960
CGACTTTTTG	ATGCACTTTC	ATCGATTATC	AAAGCAGTTG	GAACAGATGA	CAATCGAATT	4020
CGACCAATTA	CATTGGAACT	TGCATGTCTT	GTAATTCGGC	AAATTTTAAT	GAATGTTGAT	4080
GATGAAAAAG	TAAGATTACA	AATTCAAAAT	TGAGCAAAAT	CAGAATCTAA	ATTTCATAAA	4140
TTGTTTCAGT	ACATAACCAGT	TTAACGAAAT	TATGCTTCGA	AGTTCGTCTA	AACTTTTAT	4200
CATCATTTGG	ACAATATGTT	AATGGAGAGA	ATCTGTTTTT	GGAGTGGTTT	GAGGATGAAT	4260
ATGCAGAATT	TGAAGTAAGC	CAAGAGGTCC	GAAAATAATT	TAATTCATCC	TTTTTATTCA	4320
GGTGAATCAC	GTGAATTTTCG	ATATAATCGG	TCACGAAATG	CTTCTTCCTC	CAGCTGCAAC	4380
TCCTCTTTTCG	AATCTGCTAC	TTCATAAGCG	ATTGCCCAGT	GGATTTGAAG	AACGAATAAG	4440
AACTGTAGGA	AACTTTTTTAA	ATTTGAAAAT	TAATTATATA	TATATTTGCA	GCAAATCGTA	4500
TTCTACCTAC	ATATTCGAAA	ATTGGAACGA	GATTTGACCG	GTGAAGGAGA	CACAGAATTA	4560
CCTGTGAGAG	TGTTGAATTTC	TGATCAGGAA	CCAGTTGCCA	TCGGTGATTG	TATTAATTTA	4620
CGTGAGTTCA	TCTGCATAGA	AAACACCATA	TTTCTACTCA	AATTAACAAT	TTTCAGATAA	4680
TTCCGATCTT	CTATCCTGCA	CTGTGGTTCC	TCAACAATA	TGTTCTCTTG	GAAAACCTGG	4740
TGATCGTCTT	GCTCGATTCC	TTGTCACTGA	TAGACTTCAA	TTAATTCTTG	TGAAACCGGA	4800
TTCTCGAAAA	GCCGGATGGG	CAATTGTTTCG	ATTCTGTAGGA	CTTCTTCAAG	ATACAACAAT	4860
TAATGGAGAT	TCTACGGATT	CGAAAAGTTTT	GCATGTTGTG	GTGGAAGGGC	AACCTTCGAG	4920
AATTAAGGTA	AGAATACTAA	CGGGAAAAAA	AAATCAAAAA	ATTACTTCTG	TTTCAGAAAA	4980
GACATCCGGT	TTTAACTGCA	AAGTTCATAT	TCGATGATCA	CATTCCGGTG	ATGGCAGCAA	5040
AGCAACGGCT	CACCAAGGTA	ACGGAAAAAA	TAACCAAAAA	GACGGAAAGT	TATTGTAAAT	5100
GGACGAAATC	GGCGAAATTA	ATTGAAAAACG	TTTGAATTTG	CCGCTAAAAC	CAAACGAAAA	5160
CCAAACGAAA	GCGAAATTTA	ACTATCCCTT	CAGGTAGAAT	ATACATTTTA	TTTCTCTTTA	5220
TAGGGTCGCC	AAACAGCACG	TGGTCTGAAA	CTTCAGGCGA	TATGTTTCAGC	TCTTGAGATT	5280
CCACGTATCC	ATCCAGCGAC	AATGACGTCA	TCACCACGAA	TGAATCCATT	CAGAATTGTG	5340
AAAGGATGCG	CACCGGGAAG	TGTACGAAAA	ACTGTTTCCA	CATCATCATC	GTCAAGCCAA	5400
GGACGTCCCG	GACATTATTC	TGCAAACTCT	AGATCAGCAT	CTAGAAATGC	AGGAATGATA	5460
CCAGATGATC	CAACTCAACC	GAGTAGTTCT	TCGGAAAGAA	GATCCTAGGG	ATCAATATCT	5520
CTTCAGTTTC	ATCATTTTAT	GCTGTAAATT	GTATTTAAGT	ATTCTATTTC	TTTGTAGTAC	5580
TGTATTTACA	CATCGTCTAG	TTAAAATCAC	AAATCTCCGA	AAAAACAAAC	CAGTGAACAT	5640
GTGATATTTT	TCTTGCCCAT	AGTTCTCTTT	TTTTTTTGAA	ACAAAAACAA	TTACTTTTAT	5700
GCTCACCTAT	TCGAGCCATA	TTTTTTTCCC	AATTACCGGT	TGTTTTATTTT	AATTTCTTTT	5760
TTTTTTCTGT	AAATCTACTT	TATTTTTTAA	ACTGCATTTG	AGATTGTGTA	TATTTTTTCA	5820
AAATGGTTCA	AATGCCGAAT	CTATCTACTT	TTTAATCATT	ATTCAAACAG	AAAAACCGAT	5880
TATTTATTCA	GATTCTCAAA	AATGGCTGAA	AAAGCTGAAA	ATCTTCCATC	TTCTTCGGCC	5940
GAAGCTTCAG	AAGAGCCATC	ACCTCAAACT	GGACCAAATG	TGAATCAAAA	ACCATCGATT	6000
TTGGTTCTTG	GAATGGCTGG	TTCTGGAAAA	ACGACATTTG	TTCAGGTAAC	TTTCATTCAA	6060
TTTTGAGAGT	TTTCAAACAT	TACTATTTTC	AGCGTCTCAC	AGCATTCTTA	CATGCTCGTA	6120
AAACACCTCC	ATATGTGATT	AATCTGGATC	CGGCAGTTAG	CAAAGTACCT	TATCCAGTGA	6180
ATGTTGACAT	TCGAGATACT	GTGAAATACA	AGGAAGTTAT	GAAAGAATTG	GGAATGGGAC	6240
CAAATGGAGC	AATTATGACA	TGTCTTAACC	TGATGTGTAC	TCGTTTTGAT	AAAGTAATTG	6300
AGTTGATTAA	TAAGAGATCT	TCTGATTTCT	CAGTTTGTCT	TCTTGATACT	CCTGGACAAA	6360
TTGAAGCATT	CATTGGAGT	GCTAGTGGAT	CTATTATCAC	TGATTTCATTG	GCAAGTAGCC	6420
ATCCCACGGT	AAGGGATTTT	GATTTATGAA	ATCTGCTTGA	AATGAAAAAA	GATTCTAATA	6480

AATTTTTGAC	TTTTAAACAT	TTTTTACAGT	TATATTTGGT	CTATTTTCTA	TCATTA AAAAG	6540
CAAAATGAAA	AGTCGATTCT	ACTCCATATT	TATTAATTTT	GACTTTTCAG	GTGGTAATGT	6600
ACATTGTGGA	TTCCGCTCGT	GCCACAAATC	CAACTACATT	CATGTCCAAT	ATGCTCTACG	6660
CATGTTCCAT	TCTCTACCGT	ACCAAACCTT	CATTCAATTG	CGTTTTCAAC	AAAGCTGATA	6720
TTGTCAAACC	AACATTTGCA	CTCAAATGGA	TGCAAGATTT	CGAAAGATTT	GATGAAGCTT	6780
TAGAGGATGC	CAGAAGCAGT	TATATGAATG	ATTTGAGTCG	TTCAATTGAGT	CTCGTTCTTG	6840
ATGAATTCTA	TTGCGGACTG	AAAACAGGTT	TTTATTGCGA	ATAAAACCTT	TTTTAAATAA	6900
TAAATTTTCA	TTTGCGTCAG	TTCTGCAACT	GGAGAAGGAT	TCGAAGATGT	AATGACAGCA	6960
ATCGATGAAA	GTGTTGAAGC	ATACAAAAAA	GAATATGTTC	CAATGTATGA	AAAAGTGTGT	7020
GCTGAGAAAA	AACATTGGA	TGAGGAGGAG	AGAAAGAAAA	GAGATGAAGA	GGTAATTGTA	7080
GTAATTTAAT	TCTGATTATC	TTCAAATTTT	CAGACTCTGA	AAGGAAAAGC	TGTTACAGGAC	7140
CTGAACAAAG	TCGCCAATCC	CGACGAATTT	CTGGAGTCGG	AGTTGAATTC	AAAAATCGAT	7200
AGAATTCATT	TGGGCGGAGT	CGATGAAGAG	AATGAGGAGG	ATGCTGAACT	CGAAAGATCC	7260
TGATTTTCTT	TTTGTTTTGT	AATTTTTATT	CTATTTTGAT	CCCTGTTTAC	TTCTTATTGT	7320
TCTCATTTTG	TTGCGTTGTT	TTACATTTTA	CTCATTTTGT	CATAAACTTG	TTGCAAAAAT	7380
CAATATAATT	TTTGATCTGG	AAATGGTTTT	AAACCTTAAC	CTTTCATATA	TTAATAATTT	7440
TTTTTCAAAA	AAACGTTCTA	AAAAGGTTCC	TCATTTTTTT	AATATAGGAA	ATTTTGAAGA	7500
TCTTTTCCAA	AAATGAGGTT	CTTCGCTTGA	AAAGCCAACA	TTTAAACCTT	TTTTTTTTTCC	7560
AGAAACCTAG	TGGTTAATGT	CTGAAAAGAC	GTTCCACAAG	GCACAGACCA	TCCGTGCAAA	7620
GGCATCCGGA	GTGCCTTCAA	TCGTGGAAGC	TGTACAGTTT	CATGGAGTTC	GCATCACAAA	7680
AAACGATGCT	TTGGTTAAGG	AGGTACTACC	CAAATTTCAA	AATGTTGCAC	AATTCAAATTG	7740
AAAATATAAA	TTGTGAATTA	AATTCAACTT	ACATGTTTTT	TCAGGTTTCC	GAATTATACA	7800
GAAGTAAAAA	TCTAGATGAA	CTTGTTTATA	ACTCTCATCT	GGCGGCTCGT	CATCTTCAAG	7860
AAGTTGGATT	AATGGATAAT	GCAGTTGCTC	TAATTGATAC	ATCTCCAAGC	TCAAATGAAG	7920
GATATGTTGT	CAATTTCCCTA	GTTTCGAGAAC	CAAAATCATT	CACTGCTGGA	GTCAAAGCAG	7980
GAGTTTCAAC	GAATGGAGAT	GCGGATGTCA	GTTTAAATGC	CGGAAAACAA	AGTGTGGAG	8040
GACGAGGAGA	GGCAATCAAT	ACACAGTATA	CATATACTGT	AAAGGTAAGG	ACGAGAGTTG	8100
GCACTGCCAG	TTTGGCATGT	TCTCCCAATA	TTTTTTAATT	ATAAAATTTG	GAAGTATAAA	8160
AAAATGTTTG	CTTACCTATA	AAATAGCCTT	TTTACATAGA	AAAAAATTGA	AAAAAAGTGC	8220
TCAAAAATTT	CAGAAATTTT	CAATTTCCAA	ACAATTTTGG	AGAACTTTCA	AAAATTTTTC	8280
CAACTGAAAT	TAAAGCTATA	TTCTATCACT	AAATTTTATA	CAAGTCTTAA	GAGAAAATGA	8340
TGAAGTGGCT	CATTTTGTAG	AATTTCTTAA	AAAATAATAT	CTTCAGGGCG	ATCACTGCTT	8400
CAACATTTCC	GCAATCAAAC	CATTCCTGGG	ATGGCAAAAA	TATTCGAATG	TATCAGCGAC	8460
TCTATACCGT	TCACTTGCAC	ATATGCCATG	GAATCAATCA	GATGTTGATG	AGAATGCAGC	8520
TGTTCTTGCA	TATAATGGAC	AACTATGGAA	TCAAAGCTTT	TTGCATCAAG	TCAAATTGAA	8580
TGCGGTAAG	TATTATAAGT	GTTTTGTCCA	AACTATGATA	CAGTTCTTCA	GATATGGAGA	8640
ACACTTCGTG	CCACTCGAGA	TGCCGCATTT	TCAGTTCTGT	AACAAGCCGG	ACACACTTTG	8700
AAATTCTCGT	TGGAGAATGC	TGTAGCTGTT	GATACAAGAG	ATAGACCTAT	TCTTGCAAGT	8760
CGTGGAATTC	TTGGTAAGAG	TAACAACGAC	TATTTTTAAA	AAATATCTTT	TTCGAAAAAA	8820
TTACGAACGA	AAAAAACTG	TATTATGTAC	CCAAACGCGA	AATTTTGCAG	TTCTTGCGCG	8880
TTCTTGTTGA	TAAAAAATAT	GTAAAAAATT	GGAAAAACTA	CGAAAAGTCG	ATAAAAATTC	8940
CGTACCAACC	GGAAAATGTT	TCATTAATTT	CTCTTCCTTT	TTTCAGCTCG	TTTTGCTCAA	9000
GAGTACGCAG	GAGTATTTGG	TGATGCGTCA	TTTGTGAAGA	ATACATTAGA	TTTACAGGTA	9060
ACAACCTTAT	TTCAACAATT	ATTTCAAATT	CTATTA AAAA	TAATTCCAGG	CAGCTGCCCC	9120
TCTTCCACTC	GGTTTCATTC	TTGCCGCCCT	ATTCCAAGCG	AAACATTTGA	AAGGACTCGG	9180
AGATCGAGAA	GTTTCATATTT	TGGATAGATG	TTATTTGGGT	GGACAACAGG	ATGTTTCGAGG	9240
ATTTGGTCTG	AATACTATTG	GAGTGAGTTT	TAACGAAATT	CTCTTGAAAG	TCAAATAATC	9300
ATTTTCAGGT	TAAAGCAGAT	AACAGTTGTC	TTGGAGGAGG	TGCTTCACTT	GCTGGTGTCTG	9360
TTCAATTTGTA	TCGGCCATTG	ATTCCACCAA	ATATGCTATT	TGCACACGCA	TTCTTGTCAT	9420
CTGGAAGTGT	TGCATCAGTT	CATTCCAAAA	ATTTGGTGCA	ACAATTACAG	GATACTCAAC	9480
GAGTATCAGC	CGGATTTGGT	GAGTTTGAAA	TTTAGGAAAC	ATTTGGATGA	AATGTATTTT	9540
TTAAAAATAG	ATCAGCTTTA	TTTATTTGAA	AAAAAACGCT	CATTAATCAA	TAGTGATAGT	9600
TCCATTCTGA	GTTTCTTCTT	CTTCCTCGCG	GAATACAATT	TTTGACTTGT	TCGCATCCTT	9660
CTTGTTGACT	TTGTCACCAA	TCTTCTCATC	AACTAAATCT	CGAAACTGAA	AAAAATTTCAA	9720
AATTATTCCA	AAAAATATTG	ATGCAGACTA	CCTTTTTGAT	GGCTTCTGGT	ACGTTTCTAG	9780
CGTCGAATGG	ATTGGCTCCT	CCAATAATTA	AAGTCTCGTT	CGGTAGTTTA	GCCAGACGGA	9840

CGGTGTGCTT	CAACATTTTT	CTAATTAATC	TATTTCAATT	CAAGTCACTC	ACTCTCTCTT	9900
GACGTCTTCT	TCTATATTCC	AAGAACTCTG	CAGAAAATCC	GTGTCCGCCT	TGTGTGTTTC	9960
TAGTTGGCGT	CGGAGGATTC	ACGGGTCCAA	GACGAATGGA	TGCTAAAAA	ATGTTATATT	10020
TTTGATAAAA	GAAAAACACCA	TACCTTCACC	ACTTTTTGAG	TTGTGGGCGT	TCTGAATGGA	10080
ATTGATCGAT	TATTATTGCT	CTTCTTGAT	TTGCTTCTAT	CAGCTGCGTA	ATGAGGTGTT	10140
CTAAAGATCA	GCTTTAATTC	ATTTGGACAA	GTGCTCCTCT	AATAAACTTA	CCCTGTACTC	10200
ATTTTTGAAA	CGATTTACGA	TGATAAGATT	GAAAGTGGAA	GTTAAATTTA	GTCTTTCAAA	10260
GTTGAAATAA	AATCTTCATA	AATAAATAAA	TTTAAATGAA	AGATTAAATA	AATTAACGTT	10320
CACGTAGTTA	AAAAAATAAT	TTAAATCTTA	ACTTCTAATA	AAAAATCTCA	ATTTTCCAGG	10380
ACTCGCATT	GTGTTCAAAA	GTATTTTCCG	GCTGGAACCT	AACTACACGT	ATCCATTGAA	10440
ATATGTGCTC	GGCGATTCAT	TGCTCGGTGG	ATTCCATATT	GGAGCTGGTG	TCAACTTCTT	10500
GTAGAGATTA	ATTGGATGCA	AGCACCCCTC	AAAAAGATTT	TTTTGAAAAA	CGATAAATTC	10560
ACAGAATTTT	AGTTCTTTTT	CTCCCCCTTT	TATTGTTATT	TTCATCGTAA	TGCTGTGCTA	10620
GAAGTCAGAG	TAAATATGAG	TTTTTTTGTG	TTCTAGGAAT	TCCATTTTTT	CAGGAAGCAA	10680
ATTTAATAAA	AATTATCGAA	TTTCTTGCTC	TAAAGATGTT	GTACATTTTA	TGGAAATGTT	10740
CGTATAGTAA	TTCGAACACT	TTATATTTCT	CGTTTTAAAA	CTGTCGGTGT	TTTATAGTAA	10800
ACTATCTTCA	GAAAAAAATG	AGCCTACGAA	AAATCAATTT	CGTAACTGGA	AACGTGAAGA	10860
AGCTTGAAGA	AGTCAAGGCT	ATTTTGAAGA	ATTTTCGAGGT	AAAATATATT	TGATATTATT	10920
CGAACGCGAA	ATTTTGCGCC	AAAAGTACGA	TGCCTGGTCT	CAACACGACA	ATATTTTGTT	10980
AAATACAAAC	GAATGTGCGC	CTTCAAAGAA	AAGTTTCAAT	CTTTCGTTGC	CGTGGAGATA	11040
TTTTTAGAGT	TTTTGTTTAA	ATTATATATT	TGTCGTATCG	AAACCGGGTA	CCGTAATCAA	11100
TCAATTAAT	ATTTTCAGGT	TTCAAACGTG	GATGTCGATT	TGGATGAATT	CCAAGGAGAA	11160
CCCGAATTTA	TTGCCGAAAG	AAAGTGCCGT	GAGGCTGTTG	AAGCTGTAAA	AGGGCCCGTT	11220
TTGGTATGGA	AAATTGTATT	TGTTCTAAAA	ATTGTCAAAT	TTCAGGTCGA	AGACACAAGT	11280
TTATGCTTCA	ACGCAATGGG	CGGTCTTCCT	GGACCTTATA	TCAAGTGGTT	TTTGAAGAAT	11340
TTGAAACCAG	AAGGACTACA	TAATATGCTA	GGTAAATATT	TTAATTTTTT	GAAAAAACTT	11400
ATTTTTCAGC	CGGATTTTCT	GACAAAACCG	CCTATGCTCA	ATGCATCTTT	GCGTACACTG	11460
AAGGACTCGG	AAAACCTATT	CATGTATTTG	CTGGTATGAT	TTTTTGAATT	TAATTCCTTA	11520
ATTTTATATG	TTAATTTAGT	TGTTTCATTC	CTCAATTTAT	GAGAGATTTT	TTTTTCAATT	11580
TTTCTATTTT	AGGAAAATGT	CCTGGTCAAA	TTGTTGCTCC	ACGTGGTGAT	ACTGCTTTTG	11640
GATGGGATCC	ATGCTTCCAG	CCAGATGGTT	TTAAAGAAAC	ATTCGGAGAA	ATGGATAAAG	11700
ATGTAAAAAA	TGAAATTTCT	CATCGTGCAA	AGGCTCTGGA	ACTCCTCAAG	GAATATTTTC	11760
AGAATAATTA	AATTATTTTT	TCTCATCTAT	GCAATTTCTT	GAAAATTTGT	TAAGTTTCCG	11820
TTGTTATGCA	TTTGCTTTTA	TTTAAAAAAA	AAAGAATATT	TTTACATTAA	TATTAGATAT	11880
GAGAAAAGAG	TAATTTCTGG	ATTTTAACCT	TCCTACAAAA	GAATATTTAT	ATTTTTTGTA	11940
TGATTTTTTA	AAAATATCGT	CAGGAAATAA	TAACATTTCA	GATATACCTT	GAACCTTACA	12000
GTTTATGATA	TTCAGGAAAT	TTCTGAATTT	TCTGAAACCT	TACAAAATGC	GAACGGATCC	12060
GATTATTTTC	GTGATTGGGT	GCACTGGAAC	CGGGAAGATG	GATCTTGGAG	TGGCAATTGC	12120
AAAGAAATAT	GGAGGAGAGG	TGATTAGTGT	AGATTCAATG	CAATTTTATA	AAGGTACATG	12180
GGTTTTGTTT	CAATTTTAAA	TTAATTAATT	TTCTTTTTTC	AGGACTTGAC	ATTGCCACGA	12240
ATAAGATAAC	GGAAGAAGAA	TCTGAAGGGA	TTCAACATCA	TATGATGTCA	TTTTTTGAATC	12300
CATCTGAATC	ATCATCTTAT	AATGTACATA	GTTTCCGAGA	AGTCACGTTG	GATCTTATTA	12360
AAGTGCTTAA	TTCGCCACTT	TTTGAACCTG	ATCCTAATTT	TCATAATTTT	CAGAAAATCC	12420
GCGCCCGTTC	AAAAATTCCT	GTAATTGTCG	GAGGAACCAC	TTATTATGCT	GAAAGTGTC	12480
TTTATGAGAA	TAATCTGATT	GAAACCAACA	CTTCAGATGA	CGTGGATTCC	AAATCGAGAA	12540
CATCATCAGA	ATCGTCATCT	GAAGACACTG	AAGAAGGAAT	TAGTAATCAA	GAATTATGGG	12600
ATGAATTGAA	AAAAATCGAC	GAAAAATCAG	CACTTCTTCT	ACATCCAAAT	AATCGTTATC	12660
GAGTACAGAG	AGCATTGCAA	ATTTTCAGAG	AACTGGTAA	TTGATTGCA	AATTTCCAGA	12720
TTAAAAACAA	ATCAAGTAAA	GTTTTTTGCA	GGAATCCGAA	AAAGTGAAC	TGTTGAAAAA	12780
CAGAAATCAG	ATGAACTGT	TGATTTGGGT	GGACGACTAC	GATTTGATAA	TTCTTTAGTT	12840
ATTTTTATGG	ATGCAACACC	TGAAGTTTTA	GAAGAAAGAC	TTGATGGAAG	AGTTGATAAA	12900
ATGATTAAAT	TGGGTTTGAA	GAATGAATTG	ATCGAGTTTT	ATAACGAGGT	AAATATTTGA	12960
ATTTTTCCAG	AAAAAAAAG	AAAATTTTTT	ATTATTTTGT	TTTTTTTTCA	TTCTTTACTA	13020
TTTTTCAAAA	AAGTTTAAAC	TTTTGAAAAC	TGTTTCAGAAA	ATGTTTCGTG	ATTTATTTTA	13080
GCTTACTGAG	GCATTATTTT	ATTGTGATTT	TTACTATACT	CTATAAACTA	AATTTTCAGC	13140
ACGCCGAGTA	CATAAATCAC	AGCAAATATG	GTGTCATGCA	ATGTATTGGT	CTTAAAGAAT	13200

```

TCGTTCCATG GCTCAATTTG GACCCATCAG AAAGAGATAC ACTCAATGGG GATAAATTGT 13260
TCAAGCAAGG GTAATTTAAA TTTATTTTCA ATTTTATATA ATTCCAAGCT ATTTTCAGAT 13320
GCGATGATGT GAAGCTTCAC ACTCGACAAT ATGCACGGCG CCAGAGACGG TGGTATCGAT 13380
CGAGACTTTT AAAACGGTCG GATGGTGATC GGGTATGTTG ATTTTAAAAA AATTGAATTT 13440
TTAAAGAACT TTTTACTAA ATTAACAAAG TTATTGGCTG AAAATGGCTG AAAATTATAG 13500
TAAAACTAAT CAAAAAATT GAAATTTTGA ATTAAAGTCA TAAAGTGACG ACCAGAAAAT 13560
TAAAAAATAA CATTTTCTA TTTTAATTAA TTCACTCTAC TTCACTTTAA AAATAATTTT 13620
CAGAAAATGG CAAGTACAAA AATGCTGGAT ACATCTGACA AGTACCGAAT AATTAGTGAT 13680
GGAATGGACA TTGTTGATCA ATGGATGAAT GGAATCGATC TATTTGAAGA TGTAATAATTT 13740
CACAAATTCT AAAATTTCCG AATCACAAT TAAAATTTCT ACAGATCTCC ACAGACACCA 13800
ATCCAATTCT AAAAGGGTCC GATGCAAATA TTCTGCTGAA TTGTGAAATC TGTAATATTT 13860
CAATGACTGG AAAAGATAAT TGGTTTGTTC CAATACATAT TATAATTTTC AAATGAATTT 13920
TTTCAGGCAG AAACATATCG ATGGGAAAAA GCACAAGCAT CATGCTAAGC AAAAGAAATT 13980
GGCAGAGACT CGCACATAAG ACGCTATATT TATTTTTTGT TAACTTAAAT TATTTTTGTT 14040
GTTGATTGTT CTCTAAATAA AAAACAGCT CAGAGAGAAG ATTAGGCGCT CGTCCACATC 14100
TCCGACGATA GTCAACCCGA ACGAAGGGAA CTATCTTTAA TTGTCAGTGA TGACGTCATG 14160
TCGTCAAGAA CTCGTCATAG CTGTGAGAAT TGAACCATTA TAGATTTGGA CATTAGTTTA 14220
GGTTATATCC AGTACACTAA ATGGTACATG ATAGACAGTG TACATTTACA GATTTATAGA 14280
TTGTCTCAGT GACTAGTTAC CGGAAGAGGA GAGGAGAACA TGTGGCGATG TCTTTTGGAT 14340
CGATATTATT CCGTCTGAAA ATTGTTCACT AGGGGGACTG CCGATTACCA CTTCACATGA 14400
CGGAACATGT TAGTTAAAT ATTGGCTTTT ATACACATTT TCAAAATAGC ACCTGTAT 14458

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 430 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ile Phe Arg Lys Phe Leu Asn Phe Leu Lys Pro Tyr Lys Met Arg
1           5           10           15
Thr Asp Pro Ile Phe Val Ile Gly Cys Thr Gly Thr Gly Lys Ser
20           25           30
Asp Leu Gly Val Ala Ile Ala Lys Tyr Gly Gly Glu Val Ile Ser
35           40           45
Val Asp Ser Met Gln Phe Tyr Lys Gly Leu Asp Ile Ala Thr Asn Lys
50           55           60
Ile Thr Glu Glu Glu Ser Glu Gly Ile Gln His His Met Met Ser Phe
65           70           75           80
Leu Asn Pro Ser Glu Ser Ser Ser Tyr Asn Val His Ser Phe Arg Glu
85           90           95
Val Thr Leu Asp Leu Ile Lys Lys Ile Arg Ala Arg Ser Lys Ile Pro
100          105          110
Val Ile Val Gly Gly Thr Thr Tyr Tyr Ala Glu Ser Val Leu Tyr Glu
115          120          125
Asn Asn Leu Ile Glu Thr Asn Thr Ser Asp Asp Val Asp Ser Lys Ser
130          135          140
Arg Thr Ser Ser Glu Ser Ser Ser Glu Asp Thr Glu Glu Gly Ile Ser
145          150          155          160
Asn Gln Glu Leu Trp Asp Glu Leu Lys Lys Ile Asp Glu Lys Ser Ala

```

[illegible][illegible][illegible]

- [illegible]

[illegible][illegible][illegible]

```

GCATCCACAT GACAAACGCA AAGTGGCCAG GAGCTTGCAA GTTTTTGAAG AAACAGGAAT 600
CTCTCATAGT GAATTTCTCC ATCGTCAACA TACGGAAGAA GGTGGTGGTC CCCTTGGAGG 660
TCCTCTGAAG TTCTCTAACC CTTGCATCCT TTGGCTTCAT GCTGACCAGG CAGTTCTAGA 720
TGAGCGCTTG GATAAGAGGG TGGATGACAT GCTTGCTGCT GGGCTCTTGG AGGAACATAAG 780
AGATTTTTCAC AGACGCTATA ATCAGAAGAA TGTTTCGGAA AATAGCCAGG ACTATCAACA 840
TGGTATCTTC CAATCAATTG GCTTCAAGGA ATTTACAGAG TACCTGATCA CTGAGGGAAA 900
ATGCACACTG GAGACTAGTA ACCAGCTTCT AAAGAAAGGA CCTGGTCCCA TTGTCCCCC 960
TGTCTATGGC TTAGAGGTAT CTGATGTCTC GAAGTGGGAG GAGTCTGTTC TTGAACCTGC 1020
TCTTGAAATC GTGCAAAGTT TCATCCAGGG CCACAAGCCT ACAGCCACTC CAATAAAGAT 1080
GCCATACAAT GAAGCTGAGA ACAAGAGAAG TTATCACCTG TGTGACCTCT GTGATCGAAT 1140
CATCATTGGG GATCGCGAAT GGGCAGCGCA CATAAAATCC AAATCCCACT TGAACCAACT 1200
GAAGAAAAGA AGAAGATTGG ACTCAGATGC TGTCAACACC ATAGAAAGTC AGAGTGTTTC 1260
CCCAGACTAT AACAAAGAAC CTAAAGGGAA GGGATCCCCA GGGCAGAATG ATCAAGAGCT 1320
GAAATGCAGC GTTTAAGAGA CATGTCCAGT GGCCTTTGGA AAGGTGGTGG GGATCCAGTT 1380
CAGGAGGGAG GGGTATGTTT GTCTCCCAGT CTGGGCAAAG GAGTGCTATG CGGAATTCTC 1440
TGCATAGCAG AAAAGCTCCC ACCATTTTCT TTTGATGTGG TTTTAAAGTC TCACGTTCTC 1500
TATAATAGAA ACAGCAGGTC TTGTCAGCTC CTTGTGTGGC TGATGTGTCT GGAAATGATG 1560
TAGTTCAGGA AAGCATTTT TTTTCTTTG AACCTTAAAG GTTCTATTAT TAAAAGCAGC 1620
ACAGATTCCA CATTTTATA CATGAGGATC TTCTTTGTGG TGAATACCAG GATTGACTGC 1680
ATCCCTTTAA AAGAAGTTTT ATGTCCCTGA CTCTGGCTAA AATTATCTAA TTTCCAGATG 1740
CTTTTGTAGA TGAAGTGAAG ATTTGTGAGC CACATATTGG GAGTTCTAGA TTTGAGTGAA 1800
TGGCAGGAAA GGGCCATCTC CATTGAGATG ATTAAGTGAA CCAAACACTAGT TCTCGGAATT 1860
CTACAGAGAA GGAGGGAATC AGACTGAGGA AGCTGTGACA TAGGACTTGA AGACCAAAGA 1920
CTTTGAAATT TGCGAGCTGC TCATGTGTGA GTTATTATCA CTGCTGTCTT TCTATTGAGT 1980
TACAAATCTA TATTTTATT GAAGTTTAA TAAAGAAAAA ATTTACAAGA AAAAAAAAAA 2040
A 2041

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 892 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Phe Arg Lys Leu Gly Ser Ser Gly Ser Leu Trp Lys Pro Lys Asn
 1           5           10           15
Pro His Ser Leu Glu Tyr Leu Lys Tyr Leu Gln Gly Val Leu Thr Lys
          20           25           30
Asn Glu Lys Val Thr Glu Asn Asn Lys Lys Ile Leu Val Glu Ala Leu
          35           40           45
Arg Ala Ile Ala Glu Ile Leu Ile Trp Gly Asp Gln Asn Asp Ala Ser
          50           55           60
Val Phe Asp Phe Phe Leu Glu Arg Gln Met Leu Leu Tyr Phe Leu Lys
65           70           75           80
Ile Met Glu Gln Gly Asn Thr Pro Leu Asn Val Gln Leu Leu Gln Thr
          85           90           95
Leu Asn Ile Leu Phe Glu Asn Ile Arg His Glu Thr Ser Leu Tyr Phe
          100          105          110
Leu Leu Ser Asn Asn His Val Asn Ser Ile Ile Ser His Lys Phe Asp
          115          120          125

```



Leu Gln Asn Asp Glu Ile Met Ala Tyr Tyr Ile Ser Phe Leu Lys Thr  
 130 135 140  
 Leu Ser Phe Lys Leu Asn Pro Ala Thr Ile His Phe Phe Phe Asn Glu  
 145 150 155 160  
 Thr Thr Glu Glu Phe Pro Leu Leu Val Glu Val Leu Lys Leu Tyr Asn  
 165 170 175  
 Trp Asn Glu Ser Met Val Arg Ile Ala Val Arg Asn Ile Leu Leu Asn  
 180 185 190  
 Ile Val Arg Val Gln Asp Asp Ser Met Ile Ile Phe Ala Ile Lys His  
 195 200 205  
 Thr Lys Glu Tyr Leu Ser Glu Leu Ile Asp Ser Leu Val Gly Leu Ser  
 210 215 220  
 Leu Glu Met Asp Thr Phe Val Arg Ser Ala Glu Asn Val Leu Ala Asn  
 225 230 235 240  
 Arg Glu Arg Leu Arg Gly Lys Val Asp Asp Leu Ile Asp Leu Ile His  
 245 250 255  
 Tyr Ile Gly Glu Leu Leu Asp Val Glu Ala Val Ala Glu Ser Leu Ser  
 260 265 270  
 Ile Leu Val Thr Thr Arg Tyr Leu Ser Pro Leu Leu Leu Ser Ser Ile  
 275 280 285  
 Ser Pro Arg Arg Asp Asn His Ser Leu Leu Leu Thr Pro Ile Ser Ala  
 290 295 300  
 Leu Phe Phe Phe Ser Glu Phe Leu Leu Ile Val Arg His His Glu Thr  
 305 310 315 320  
 Ile Tyr Thr Phe Leu Ser Ser Phe Leu Phe Asp Thr Gln Asn Thr Leu  
 325 330 335  
 Thr Thr His Trp Ile Arg His Asn Glu Lys Tyr Cys Leu Glu Pro Ile  
 340 345 350  
 Thr Leu Ser Ser Pro Thr Gly Glu Tyr Val Asn Glu Asp His Val Phe  
 355 360 365  
 Phe Asp Phe Leu Leu Glu Ala Phe Asp Ser Ser Gln Ala Asp Asp Ser  
 370 375 380  
 Lys Ala Phe Tyr Gly Leu Met Leu Ile Tyr Ser Met Phe Gln Asn Asn  
 385 390 395 400  
 Ala Asp Val Gly Glu Leu Leu Ser Ala Ala Asn Phe Pro Val Leu Lys  
 405 410 415  
 Glu Ser Thr Thr Thr Ser Leu Ala Gln Gln Asn Leu Ala Arg Leu Arg  
 420 425 430  
 Ile Ala Ser Thr Ser Ser Ile Ser Lys Arg Thr Arg Ala Ile Thr Glu  
 435 440 445  
 Ile Gly Val Glu Ala Thr Glu Glu Asp Glu Ile Phe His Asp Val Pro  
 450 455 460  
 Glu Glu Gln Thr Leu Glu Asp Leu Val Asp Asp Val Leu Val Asp Thr  
 465 470 475 480  
 Glu Asn Ser Ala Ile Ser Asp Pro Glu Pro Lys Asn Val Glu Ser Glu  
 485 490 495  
 Ser Arg Ser Arg Phe Gln Ser Ala Val Asp Glu Leu Pro Pro Pro Ser  
 500 505 510  
 Thr Ser Gly Cys Asp Gly Arg Leu Phe Asp Ala Leu Ser Ser Ile Ile  
 515 520 525  
 Lys Ala Val Gly Thr Asp Asp Asn Arg Ile Arg Pro Ile Thr Leu Glu  
 530 535 540  
 Leu Ala Cys Leu Val Ile Arg Gln Ile Leu Met Thr Val Asp Asp Glu  
 545 550 555 560  
 Lys Val His Thr Ser Leu Thr Lys Leu Cys Phe Glu Val Arg Leu Lys  
 565 570 575

Leu	Leu	Ser	Ser	Ile	Gly	Gln	Tyr	Val	Asn	Gly	Glu	Asn	Leu	Phe	Leu
			580					585					590		
Glu	Trp	Phe	Glu	Asp	Glu	Tyr	Ala	Glu	Phe	Glu	Val	Asn	His	Val	Asn
		595					600					605			
Phe	Asp	Ile	Ile	Gly	His	Glu	Met	Leu	Leu	Pro	Pro	Ala	Ala	Thr	Pro
	610					615					620				
Leu	Ser	Asn	Leu	Leu	Leu	His	Lys	Arg	Leu	Pro	Ser	Gly	Phe	Glu	Glu
625					630					635					640
Arg	Ile	Arg	Thr	Gln	Ile	Val	Phe	Tyr	Leu	His	Ile	Arg	Lys	Leu	Glu
			645						650					655	
Arg	Asp	Leu	Thr	Gly	Glu	Gly	Asp	Thr	Glu	Leu	Pro	Val	Arg	Val	Leu
		660					665						670		
Asn	Ser	Asp	Gln	Glu	Pro	Val	Ala	Ile	Gly	Asp	Cys	Ile	Asn	Leu	His
		675					680					685			
Asn	Ser	Asp	Leu	Leu	Ser	Cys	Thr	Val	Val	Pro	Gln	Gln	Leu	Cys	Ser
	690					695					700				
Leu	Gly	Lys	Pro	Gly	Asp	Arg	Leu	Ala	Arg	Phe	Leu	Val	Thr	Asp	Arg
705				710						715					720
Leu	Gln	Leu	Ile	Leu	Val	Glu	Pro	Asp	Ser	Arg	Lys	Ala	Gly	Trp	Ala
			725						730					735	
Ile	Val	Arg	Phe	Val	Gly	Leu	Leu	Gln	Asp	Thr	Thr	Ile	Asn	Gly	Asp
		740						745					750		
Ser	Thr	Asp	Ser	Lys	Val	Leu	His	Val	Val	Val	Glu	Gly	Gln	Pro	Ser
		755					760					765			
Arg	Ile	Lys	Lys	Arg	His	Pro	Val	Leu	Thr	Ala	Lys	Phe	Ile	Phe	Asp
	770					775					780				
Asp	His	Ile	Arg	Cys	Met	Ala	Ala	Lys	Gln	Arg	Leu	Thr	Lys	Gly	Arg
785				790						795					800
Gln	Thr	Ala	Arg	Gly	Leu	Lys	Leu	Gln	Ala	Ile	Cys	Ser	Ala	Leu	Gly
			805						810					815	
Val	Pro	Arg	Ile	Asp	Pro	Ala	Thr	Met	Thr	Ser	Ser	Pro	Arg	Met	Asn
		820						825					830		
Pro	Phe	Arg	Ile	Val	Lys	Gly	Cys	Ala	Pro	Gly	Ser	Val	Arg	Lys	Thr
	835					840						845			
Val	Ser	Thr	Ser	Ser	Ser	Ser	Ser	Gln	Gly	Arg	Pro	Gly	His	Tyr	Ser
	850					855					860				
Ala	Asn	Leu	Arg	Ser	Ala	Ser	Arg	Asn	Ala	Gly	Met	Ile	Pro	Asp	Asp
865				870						875					880
Pro	Thr	Gln	Pro	Ser	Ser	Ser	Ser	Glu	Arg	Arg	Ser				
			885						890						

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 355 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ala	Glu	Lys	Ala	Glu	Asn	Leu	Pro	Ser	Ser	Ser	Ala	Glu	Ala	Ser
1				5				10					15		

Glu	Glu	Pro	Ser	Pro	Gln	Thr	Gly	Pro	Asn	Val	Asn	Gln	Lys	Pro	Ser
			20					25					30		
Ile	Leu	Val	Leu	Gly	Met	Ala	Gly	Ser	Gly	Lys	Thr	Thr	Phe	Val	Gln
		35					40					45			
Arg	Leu	Thr	Ala	Phe	Leu	His	Ala	Arg	Lys	Thr	Pro	Pro	Tyr	Val	Ile
	50					55					60				
Asn	Leu	Asp	Pro	Ala	Val	Ser	Lys	Val	Pro	Tyr	Pro	Val	Asn	Val	Asp
65					70					75				80	
Ile	Arg	Asp	Thr	Val	Lys	Tyr	Lys	Glu	Val	Met	Lys	Glu	Phe	Gly	Met
				85				90						95	
Gly	Pro	Asn	Gly	Ala	Ile	Met	Thr	Cys	Leu	Asn	Leu	Met	Cys	Thr	Arg
			100					105					110		
Phe	Asp	Lys	Val	Ile	Glu	Leu	Ile	Asn	Lys	Arg	Ser	Ser	Asp	Phe	Ser
		115					120					125			
Val	Cys	Leu	Leu	Asp	Thr	Pro	Gly	Gln	Ile	Glu	Ala	Phe	Thr	Trp	Ser
	130					135					140				
Ala	Ser	Gly	Ser	Ile	Ile	Thr	Asp	Ser	Leu	Ala	Ser	Ser	His	Pro	Thr
145					150					155					160
Val	Val	Met	Tyr	Ile	Val	Asp	Ser	Ala	Arg	Ala	Thr	Asn	Pro	Thr	Thr
				165				170						175	
Phe	Met	Ser	Asn	Met	Leu	Tyr	Ala	Cys	Ser	Ile	Leu	Tyr	Arg	Thr	Lys
			180					185					190		
Leu	Pro	Phe	Ile	Val	Val	Phe	Asn	Lys	Ala	Asp	Ile	Val	Lys	Pro	Thr
		195					200					205			
Phe	Ala	Leu	Lys	Trp	Met	Gln	Asp	Phe	Glu	Arg	Phe	Asp	Glu	Ala	Leu
	210					215					220				
Glu	Asp	Ala	Arg	Ser	Ser	Tyr	Met	Asn	Asp	Leu	Ser	Arg	Ser	Leu	Ser
225					230					235					240
Leu	Val	Leu	Asp	Glu	Phe	Tyr	Cys	Gly	Leu	Lys	Thr	Val	Cys	Val	Ser
				245				250						255	
Ser	Ala	Thr	Gly	Glu	Gly	Phe	Glu	Asp	Val	Met	Thr	Ala	Ile	Asp	Glu
			260					265					270		
Ser	Val	Glu	Ala	Tyr	Lys	Lys	Glu	Tyr	Val	Pro	Met	Tyr	Glu	Lys	Val
		275					280					285			
Leu	Ala	Glu	Lys	Lys	Leu	Leu	Asp	Glu	Glu	Glu	Arg	Lys	Lys	Arg	Asp
	290					295					300				
Glu	Glu	Thr	Leu	Lys	Gly	Lys	Ala	Val	His	Asp	Leu	Asn	Lys	Val	Ala
305					310					315					320
Asn	Pro	Asp	Glu	Phe	Leu	Glu	Ser	Glu	Leu	Asn	Ser	Lys	Ile	Asp	Arg
				325				330						335	
Ile	His	Leu	Gly	Gly	Val	Asp	Glu	Glu	Asn	Glu	Glu	Asp	Ala	Glu	Leu
			340					345					350		
Glu	Arg	Ser													
		355													

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Ser	Glu	Lys	Thr	Phe	His	Lys	Ala	Gln	Thr	Ile	Arg	Ala	Lys	Ala	1	5	10	15
Ser	Gly	Val	Pro	Ser	Ile	Val	Glu	Ala	Val	Gln	Phe	His	Gly	Val	Arg	20	25	30	
Ile	Thr	Lys	Asn	Asp	Ala	Leu	Val	Lys	Glu	Val	Ser	Glu	Leu	Tyr	Arg	35	40	45	
Ser	Lys	Asn	Leu	Asp	Glu	Leu	Val	His	Asn	Ser	His	Leu	Ala	Ala	Arg	50	55	60	
His	Leu	Gln	Glu	Val	Gly	Leu	Met	Asp	Asn	Ala	Val	Ala	Leu	Ile	Asp	65	70	75	80
Thr	Ser	Pro	Ser	Ser	Asn	Glu	Gly	Tyr	Val	Val	Asn	Phe	Leu	Val	Arg	85	90	95	
Glu	Pro	Lys	Ser	Phe	Thr	Ala	Gly	Val	Lys	Ala	Gly	Val	Ser	Thr	Asn	100	105	110	
Gly	Asp	Ala	Asp	Val	Ser	Leu	Asn	Ala	Gly	Lys	Gln	Ser	Val	Gly	Gly	115	120	125	
Arg	Gly	Glu	Ala	Ile	Asn	Thr	Gln	Tyr	Thr	Tyr	Thr	Val	Lys	Gly	Asp	130	135	140	
His	Cys	Phe	Asn	Ile	Ser	Ala	Ile	Lys	Pro	Phe	Leu	Gly	Trp	Gln	Lys	145	150	155	160
Tyr	Ser	Asn	Val	Ser	Ala	Thr	Leu	Tyr	Arg	Ser	Leu	Ala	His	Met	Pro	165	170	175	
Trp	Asn	Gln	Ser	Asp	Val	Asp	Glu	Asn	Ala	Ala	Val	Leu	Ala	Tyr	Asn	180	185	190	
Gly	Gln	Leu	Trp	Asn	Gln	Lys	Leu	Leu	His	Gln	Val	Lys	Leu	Asn	Ala	195	200	205	
Ile	Trp	Arg	Thr	Leu	Arg	Ala	Thr	Arg	Asp	Ala	Ala	Phe	Ser	Val	Arg	210	215	220	
Glu	Gln	Ala	Gly	His	Thr	Leu	Lys	Phe	Ser	Leu	Glu	Asn	Ala	Val	Ala	225	230	235	240
Val	Asp	Thr	Arg	Asp	Arg	Pro	Ile	Leu	Ala	Ser	Arg	Gly	Ile	Leu	Ala	245	250	255	
Arg	Phe	Ala	Gln	Glu	Tyr	Ala	Gly	Val	Phe	Gly	Asp	Ala	Ser	Phe	Val	260	265	270	
Lys	Asn	Thr	Leu	Asp	Leu	Gln	Ala	Ala	Ala	Pro	Leu	Pro	Leu	Gly	Phe	275	280	285	
Ile	Leu	Ala	Ala	Ser	Phe	Gln	Ala	Lys	His	Leu	Lys	Gly	Leu	Gly	Asp	290	295	300	
Arg	Glu	Val	His	Ile	Leu	Asp	Arg	Cys	Tyr	Leu	Gly	Gly	Gln	Gln	Asp	305	310	315	320
Val	Arg	Gly	Phe	Gly	Leu	Asn	Thr	Ile	Gly	Val	Lys	Ala	Asp	Asn	Ser	325	330	335	
Cys	Leu	Gly	Gly	Gly	Ala	Ser	Leu	Ala	Gly	Val	Val	His	Leu	Tyr	Arg	340	345	350	
Pro	Leu	Ile	Pro	Pro	Asn	Met	Leu	Phe	Ala	His	Ala	Phe	Leu	Ala	Ser	355	360	365	
Gly	Ser	Val	Ala	Ser	Val	His	Ser	Lys	Asn	Leu	Val	Gln	Gln	Leu	Gln	370	375	380	
Asp	Thr	Gln	Arg	Val	Ser	Ala	Gly	Phe	Gly	Leu	Ala	Phe	Val	Phe	Lys	385	390	395	400
Ser	Ile	Phe	Arg	Leu	Glu	Leu	Asn	Tyr	Thr	Tyr	Pro	Leu	Lys	Tyr	Val	405	410	415	
Leu	Gly	Asp	Ser	Leu	Leu	Gly	Gly	Phe	His	Ile	Gly	Ala	Gly	Val	Asn	420	425	430	

Phe Leu

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

Met Leu Tyr Ile Leu Trp Lys Leu Asn Tyr Leu Gln Lys Lys Met Ser
 1           5           10           15
Leu Arg Lys Ile Asn Phe Val Thr Gly Asn Val Lys Lys Leu Glu Glu
 20           25           30
Val Lys Ala Ile Leu Lys Asn Phe Glu Val Ser Asn Val Asp Val Asp
 35           40           45
Leu Asp Glu Phe Gln Gly Glu Pro Glu Phe Ile Ala Glu Arg Lys Cys
 50           55           60
Arg Glu Ala Val Glu Ala Val Lys Gly Pro Val Leu Val Glu Asp Thr
 65           70           75           80
Ser Leu Cys Phe Asn Ala Met Gly Gly Leu Pro Gly Pro Tyr Ile Lys
 85           90           95
Trp Phe Leu Lys Asn Leu Lys Pro Glu Gly Leu His Asn Met Leu Ala
 100          105          110
Gly Phe Ser Asp Lys Thr Ala Tyr Ala Gln Cys Ile Phe Ala Tyr Thr
 115          120          125
Glu Gly Leu Gly Lys Pro Ile His Val Phe Ala Gly Lys Cys Pro Gly
 130          135          140
Gln Ile Val Ala Pro Arg Gly Asp Thr Ala Phe Gly Trp Asp Pro Cys
 145          150          155          160
Phe Gln Pro Asp Gly Phe Lys Glu Thr Phe Gly Glu Met Asp Lys Asp
 165          170          175
Val Lys Asn Glu Ile Ser His Arg Ala Lys Ala Leu Glu Leu Leu Lys
 180          185          190
Glu Tyr Phe Gln Asn Asn
 195

```

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGAACACTTT ATATTCTCG

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GATAGTTCCC TTCGTTCTGGG

20

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTTCTGGATT TTAACCTTCC

20

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTCCGAGAA GTCACGTTGG

20

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TACAGGAATT TTTGAACGGG

20

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CTTCAGATGA CGTGGATTCC

20

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGAATCCGAA AAAGTGAAC

20

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AAGAGATACA CTCAATGGGG

20

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATCGATACCA CCGTCTCTGG

20

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TTGAATCTAC ACTAATCACC

20

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAATTATCT TTTCCAGTCA

20

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACATTATAAA GTTACTGTCC

20

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TTTTAGTTAA AGCATTGACC

20



## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACATCTTTAT CCATTCTCTCC

20

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGCAAAGGCT CTGGAAGCTCC

20

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

AAAAACCACT TGATATAAGG

20

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CATCCAAAAG CAGTATCACC

20

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTAATTGGAT GCAAGCACCC C

21

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATTACTATAC GAACATTTCC

20

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TTGTAAAGGC GTTAGTTTGG

20

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CAGGAGTATT TGGTGATGCG

20

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGACGGGGAG AAGGTGACGG

20

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AAAACCTTCTA CCAACAATGG

20

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGTAATCTCT CTCGATTAGC

20

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CCGTGGGATG GCTACTTGCC

20

## (2) INFORMATION FOR SEQ ID NO:33:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TGGATTTGTG GCACGAGCGG

20

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

TTGATTGCCT CTCCTCGTCC

20

## (2) INFORMATION FOR SEQ ID NO:35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

ATCAACATCT GATTGATTCC

20

## (2) INFORMATION FOR SEQ ID NO:36:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CAGCGAGCGC ATGCAACTAT ATATTGAGCA GG

32

## (2) INFORMATION FOR SEQ ID NO:37:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AATAAATATT TAAATATTCA GATATACCCT GAACTCTACA G

41

## (2) INFORMATION FOR SEQ ID NO:38:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 45 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AAACTGTAGA GTTCAGGGTA TATCTGAATA TTAAATATT TATTC

45

## (2) INFORMATION FOR SEQ ID NO:39:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTACGTGGAG CTCTGCAACT ATATATTGAG CAGG

34

## (2) INFORMATION FOR SEQ ID NO:40:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGACACTGC AGGATAGTTC CCTTCGTTTC GG

32

## (2) INFORMATION FOR SEQ ID NO:41:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GTGTTGCATC AGTTCATTCC

20

## (2) INFORMATION FOR SEQ ID NO:42:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GCTGTGCTAG AAGTCAGAGG

20

## (2) INFORMATION FOR SEQ ID NO:43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GTTCTCCTTG GAATTCATCC

20

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

AGTATATCTA GATGTGCGAG TCTTGCCAA TT

32

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

AGTAATTGTA CATTAGTGG

20

## (2) INFORMATION FOR SEQ ID NO:46:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

ATTAACCTTA CTTACTTACC

20

## (2) INFORMATION FOR SEQ ID NO:47:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CTAAACTAAG TAATATAACC

20

## (2) INFORMATION FOR SEQ ID NO:48:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GTTGATTCTT TGAGCACTGG

20

## (2) INFORMATION FOR SEQ ID NO:49:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AATTCGACCA ATTACATTGG

20

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AACATAGTTG TTGAGGAAGG

20

## (2) INFORMATION FOR SEQ ID NO:51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

AATTAATGGA GATTCTACGG

20

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TCAGCATCTA GAAATGCAGG

20



## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CGAATGTCAA CATTCACTGG

20

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

CTTAACCTGA TGTGTACTCG

20

## (2) INFORMATION FOR SEQ ID NO:55:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

ATGAAGCTTT AGAGGATGCC

20

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CGACGAATTT CTGGAGTCGG

20

## (2) INFORMATION FOR SEQ ID NO:57:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

ACTGCATTAT CCATTAATCC

20

## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

CACCCAAATA ACATCTATCC

20

## (2) INFORMATION FOR SEQ ID NO:59:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTTAACCTCA TCTTCGCTGG

20

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

ATGTTCCGCA AGCTTGGTTC

20

## (2) INFORMATION FOR SEQ ID NO:61:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TTTAATTACC CAAGTTTGAG

20

## (2) INFORMATION FOR SEQ ID NO:62:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

TTTTAACCCA GTTACTCAAG

20